JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **159**, 27–43 (1991)

Numerical Methods for Nonlinear Optimal Control Problems: Application to Abnormal Problems

ROHAN DAPATADU, CHARLES GIBSON, AND JOHN GREGORY*

Department of Mathematics, Southern Illinois University, Carbondale, IL 62901-4408

Submitted by V. Lakshmikantham

Received April 27, 1989

New, efficient, accurate numerical methods are given for general classes of optimal control problems. It is seen that these results satisfy an a priori maximum pointwise component error of $O(h^2)$ with a Richardson error of $O(h^4)$. The foundation for these results is the corresponding results for the m-dependent-variable problem in the calculus of variations and the use of multipliers to convert the optimal control problems to the calculus of variations setting. © 1991 Academic Press, Inc.

I. INTRODUCTION

New numerical methods are given for nonlinear optimal control problems. By this we mean problems with an objective functional, a trajectory equation, free or fixed boundary conditions, and equality and inequality constraints. These problems include the class of linear regulator problems as a special example (see [1, 15]). The methods of this paper "immediately" lead to higher order algorithms, the solution of initial value problems with the same accuracy, more generalized transversality conditions, and methods for minimum time problems.

Of special interest is the efficiency and accuracy of our algorithms. We establish, for the first time in this general problem setting, pointwise a priori error estimates with maximum error at the node point, $\|e\|_\infty$, equal to $O(h^2)$ and a Richardson error of $O(h^4)$. This is done under the weak assumption that there are no conjugate points for a related problem and not the usual convexity assumptions.

Of practical interest is that these methods (i) are easy to implement, (ii) hold for well defined mixtures of initial value and boundary value problems, (iii) use multipliers, and not ill-conditioned penalty methods, for

* Work supported in part by Office of Naval Research Grant N00014-88-K-0081.

0022-247X/91 $3.00

both equality and inequality constraints in a natural, efficient manner, and (iv) require the solution of nonlinear block tridiagonal systems.

The remainder of this paper is as follows. In Section 2 we define our basic problem without constraints and show that it reduces to a well defined equivalent problem in the calculus of variations. In Section 3 we give an algorithm for the calculus of variations problem and the a priori error estimates. In Section 4 we consider linear regulator type problems as a special case. In Section 5 we discuss how to handle simple equality and inequality constraints. In Section 6 we indicate how to handle simpler problems where u may be solved, at least in part, in terms of x and x'. In Section 7 we present some numerical examples to justify our theory. In Section 8 we show that our methods are sufficiently general to handle "abnormal problems" where conjugate points are generalized to focal interval with extremal solutions which vanish identically on nontrivial subintervals of $[a, b]$.

II. Basic Problems

The purpose of this section is to consider basic optimal control problems and show that they can be put into an equivalent calculus of variations problem setting.

Our basic problem is to minimize the functional

$$J_1(x) = \int_a^b f(t, x, u)\, dt \tag{1}$$

subject to the trajectory equation

$$x'(t) = g(t, x, u) \tag{2}$$

and the boundary conditions

$$x(a) = x_a, \qquad x(b) = x_b \qquad \text{or} \tag{3a}$$

$$x(a) = x_a, \qquad x(b) \text{ arbitrary}. \tag{3b}$$

Condition (3a) is the two point boundary value problem, condition (3b) corresponds to the classical linear regulator conditions.

Throughout this paper we assume that (1)–(3) has a unique solution (x^*, u^*) and enough smoothness on f and g to yield our results, below. In addition, we assume that $x(t) = (x^1(t), ..., x^n(t))^T$ is an n-vector, $u(t) = (u^1(t), ..., u^m(t))$ is an m-vector, and f_{uu} is invertible. Problems where F_{uu} is not invertible, such as minimal time problems, will be handled in later papers by the third author.

We define

$$x_1(t) = x(t) = (x^1(t), x^2(t), ..., x^n(t))^T \tag{4a}$$

$$x_2(t) = \int_a^t u(s)\, ds, \qquad x_2(a) = 0 \tag{4b}$$

and assume that x_1, x_2 are piecewise smooth and continuous and that $x_2'(t) = u(t)$ except at discontinuities of u.

THEOREM 1. *(x^*, u^*) is the unique solution to (1)–(3) iff (x_1^*, x_2^*) is the unique solution to the problem of minimizing*

$$J_2(x_1, x_2) = \int_a^b f(t, x_1, x_2')\, dt \tag{5}$$

subject to (3) *and*

$$x_1' - g(t, x_1, x_2') = 0. \tag{6}$$

To prove this result we define $\lambda(t)$ to be the Lagrange multiplier for the above problem; i.e., we consider the problem of finding a stationary point of

$$J_3(x_1, x_2, \lambda) = J_2(x_1, x_2) + \int_a^b \lambda^T[x_1' - g(t, x_1, x_2')]\, dt \tag{7}$$

subject to (3). We define $X = (x_1, x_2, \lambda)$ as the $(n + m + n = 2n + m)$-dependent-variable vector and $Y = (y_1, y_2, \mu)$ as the associated variation. Letting $H(\varepsilon) = J_3(X + \varepsilon Y)$, the first necessary condition is that

$$\begin{aligned} 0 = H'(0) &= \int_a^b \{ y_1^T f_{x_1} + y_2'^T f_{x_2'} + \mu^T[x_1' - g(t, x_1, x_2')] \\ &\quad + \lambda^T[y_1' - g_{x_1} y_1 - g_{x_2'} y_2'] \}\, dt \\ &= \int_a^b [y_1^T(f_{x_1} - g_{x_1}^T \lambda) + y_1'^T \lambda]\, dt + \int_a^b y_2'^T(f_{x_2'} - g_{x_2'}^T \lambda)\, dt \\ &\quad + \int_a^b \mu^T[x_1' - g(t, x_1, x_2')]\, dt, \end{aligned}$$

where f_{x_1} is an n-vector evaluated along the extremal solution. Similar comments hold for the n-vector g, the $n \times n$ matrix g_{x_1}, the $n \times m$ matrix $g_{x_2'}$, the m-vector $f_{x_2'}$ or the $m \times m$ matrix f_{uu}. Thus, between corners we have by integration by parts

$$\frac{d}{dt}\begin{bmatrix} \lambda \\ f_{x_2'} - g_{x_2'}^T \lambda \end{bmatrix} = \begin{bmatrix} f_{x_1} - g_{x_1}^T \lambda \\ 0 \end{bmatrix} \tag{8a}$$

$$0 = y_1^T \lambda + y_2^T (f_{x_2'} - g_{x_2'}^T \lambda) \,|_a^b \qquad \text{and} \tag{8b}$$

$$x_1' = g(t, x_1 x_2'). \tag{8c}$$

We show that these conditions are equivalent to the usual solution.

The usual solution for (1)–(3) is given by Pontryagin's maximum principle (see [1 or 12]). We define the Hamiltonian H for this problem by

$$H = f(t, x, u) + p^T g(t, x, u), \tag{9a}$$

where $p(t)$ has the same dimension and smoothness as $x(t)$. The necessary conditions are that the optimal pair (x^*, u^*) along with the optimal multiplier p^* satisfy

$$x' = H_p = g(t, x, u) \tag{9b}$$

$$p' = -H_x = -f_x - p^T g_x \tag{9c}$$

$$0 = H_u = f_u + p^T g_u \tag{9d}$$

$$p(b) = 0 \qquad \text{if} \quad x(b) \text{ is arbitrary.} \tag{9e}$$

Since we assume that f_{uu} is invertible, (9d) can be solved for u explicitly in practice so that the modified (9b)–(9c) equations are a system of $2n$ first order differential equations with $2n$ boundary conditions, which can be solved uniquely. We will see below that the conditions are (3a) or $x(a) = 0 = p(b)$ if (3b) holds.

It is a straightforward exercise to see that if we set $p(t) = -\lambda(t)$ then the conditions (8) and (9) are equivalent. The only difficult part is the transversality condition and making sense of (8b).

If (3a) holds we have along with (4b) that $y_1(a) = y_1(b) = y_2(a) = 0$ and $y_2(b)$ is arbitrary, which implies using (8b) that the function $q(t) = (f_{x_2'} - g_{x_2'}^T \lambda)(t)$ satisfies $q(b) = 0$. From the second column of (8a) we have that $q(t)$ is identically constant so that (9d) holds. If (3b) holds along with (4b) we have that $y_1(a) = y_2(a) = 0$ and $y_1(b)$ and $y_2(b)$ are arbitrary. Thus, as before, (9d) holds since $y_2(b)$ is arbitrary. In addition, that $y_1(b)$ is arbitrary implies from (8b) that $\lambda(b) = 0$ and hence $p(b) = 0$.

To complete our transformation to a problem in the calculus of variations we define $x_3(t)$ by

$$x_3(t) = \int_a^t \lambda(s)\, ds, \qquad x_3(a) = 0, \tag{10}$$

and the $(2n + m)$-dimensional vector $X(t)$ by

$$X^T(t) = (x_1^T(t), x_2^T(t), x_3^T(t)), \tag{11a}$$

where the components are given by

$$(X_1, X_2, ..., X_n)^T = x_1(t), \qquad (X_{n+1}, ..., X_{n+m})^T = x_2(t)$$

and (11b)

$$(X_{n+m+1}, ..., X_{2n+m})^T = x_3(t).$$

We define the real valued function $F(t, X, X')$ by

$$\begin{aligned} F(t, X, X') &= f(t, x_1, x_2') + x_3'^T[x_1' - g(t, x_1, x_2')] \\ &= f(t, X_1, ..., X_n, X'_{n+1}, ..., X'_{n+m}) \\ &\quad + \begin{pmatrix} X'_{2n+1} \\ \vdots \\ X'_{2n-m} \end{pmatrix}^T \left(\begin{pmatrix} X'_1 \\ \vdots \\ X'_n \end{pmatrix} \right. \\ &\quad \left. - \begin{pmatrix} g^1(t, X_1, ..., X_n, X'_{n+1}, ..., X'_{n+m}) \\ \vdots \\ g^n(t, X_1, ..., X_n, X'_{n+1}, ..., X'_{n+m}) \end{pmatrix} \right) \end{aligned} \tag{12}$$

and note that

$$F_X = \begin{array}{c} \begin{pmatrix} f_{x_1} - g_{x_1}^T x_3' \\ 0 \\ 0 \end{pmatrix} \begin{array}{l} n \\ m \\ n \end{array} \end{array} = \begin{array}{c} \begin{pmatrix} \begin{pmatrix} f_{X_1} \\ \vdots \\ f_{X_n} \end{pmatrix} - \begin{pmatrix} g^1_{X_1} & g^2_{X_1} & g^n_{X_1} \\ & \vdots & \\ g^1_{X_n} & g^2_{X_n} & g^n_{X_n} \end{pmatrix} \begin{pmatrix} X'_{n+m+1} \\ \vdots \\ X'_{2n+m} \end{pmatrix} \\ 0 \\ 0 \end{pmatrix} \begin{array}{l} n \\ m \\ n \end{array} \end{array} \tag{13a}$$

and

$$F_{X'} = \begin{pmatrix} x_3' \\ f_{x_2'} - g_{x_2'}^T x_3' \\ x_1' - g \end{pmatrix} \begin{array}{l} n \\ m \\ n \end{array}$$

$$\begin{pmatrix} \begin{pmatrix} X'_{n+m+1} \\ \vdots \\ X'_{2n+m} \end{pmatrix} \\ \begin{pmatrix} f_{X'_{n+1}} \\ \vdots \\ f_{X'_{n+m}} \end{pmatrix} - \begin{pmatrix} g^1_{X'_{n+1}} & g^2_{X'_{n+1}} & g^n_{X'_{n+1}} \\ & \vdots & \\ g^1_{X'_{n+m}} & g^2_{X'_{n+m}} & g^n_{X'_{n+m}} \end{pmatrix} \begin{pmatrix} X'_{n+m+1} \\ \vdots \\ X'_{2n+m} \end{pmatrix} \\ \begin{pmatrix} X'_1 \\ \vdots \\ X'_n \end{pmatrix} - \begin{pmatrix} g^1 \\ \vdots \\ g^n \end{pmatrix} \end{pmatrix} \begin{array}{l} n \\ m \\ n \end{array} \tag{13b}$$

For completeness we also note that

$$F_{X'X'} = \begin{array}{c} \begin{pmatrix} 0 & 0 & I \\ 0 & f_{x_2'x_2'} - g_{x_2'x_2'}^T x_3' & -g_{x_2'} \\ I & -g_{x_2'} & 0 \end{pmatrix} \end{array} \begin{array}{l} n \\ m \\ n \end{array}$$

is invertible if $f_{x_2'x_2'} - x_3'^T g_{x_2'x_2'}$ is.

We assume that $f_{x_2'x_2'} - x_3' g_{x_2'x_2'}$ has rank m evaluated along the solution so that $F_{X'X'}$ is invertible. Thus,

THEOREM 2. *(x^*, u^*) is the unique solution to (1)–(3) iff X^* gives the unique minimum to*

$$J_4(X) = \int_a^b F(t, X, X')\,dt. \tag{14}$$

In the above, X^* is defined so that $X^* = (x_1^{*T}, x_2^{*T}, x_3^{*T})^T$, where x_1^*, x_2^*, and x_3^* are given, respectively, in (4) and (10). We note that natural boundary conditions or transversality conditions still must be determined. This will be done in the next section.

To anticipate the results of the next section we note that if we wish to minimize (14), where $X(a)$ and $X(b)$ are specified, we have that the first variation $J_4(X, Y)$ satisfies

$$\begin{aligned} 0 = J_4(X, Y) &= \int_a^b \{ y_1^T f_{x_1} + y_2'^T f_{x_2} + y_3'^T [x_1' - g(t, x_1, x_2')] \\ &\quad + x_3' [y_1' - g_{x_1} y_1 - g_{x_2'} y_2'] \}\,dt \\ &= \int_a^b \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}'^T \begin{pmatrix} x_3' \\ f_{x_2} - g_{x_2}^T x_3' \\ x_1' - g(t, x_1, x_2') \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} f_{x_1} - g_{x_1}^T x_3' \\ 0 \\ 0 \end{pmatrix} \right\} dt \end{aligned}$$

or

$$0 = J_4(X, Y) = \int_a^b [Y'^T(t)\, F_{X'} + Y^T(t)\, F_X]\,dt, \tag{15}$$

where $F_{X'}$ and F_X are given in (13).

III. NUMERICAL METHODS

The purpose of this section is to show that our basic problem (1)–(3) can be solved efficiently and accurately. This will be done by converting the problem (1)–(3) into (14) and using algorithm (16) and Theorem 3, below. The results for the calculus of variations problem given in (14) are found in [9].

For ease of presentation we assume, until after Theorem 3 below, a general F in (14) for a specific two point boundary value problem where $X(a)=X_a$ and $X(b)=X_b$, with a general first variation given in (15). We assume also that X is an M-vector. After Theorem 3 we consider our specific problem with F given in (12) and $M=2n+m$.

Thus, let $\pi=(a=a_0<a_1<\cdots<a_N=b)$ be a partition of $[a,b]$ with $a_{k+1}-a_k=h=(b-a)/N$; $z_k(t)$ be the spline hat functions with $z_k(a_k)=1$, $Z_k(t)=z_k(t)\,I_{M\times M}$

$$x_h(t)=\sum_{k=0}^{N} Z_k(t)\,C_k \qquad \text{and} \qquad Y_h(t)=\sum_{k=0}^{N} Z_k(t)\,D_k$$

be, respectively, the numerical solution to our problem and the numerical admissible variation; and utilizing the linearity of $Y(t)$ in (15) we have the algorithm

$$\begin{aligned} &F_{X'}\left(a^*_{k-1},\frac{X_k+X_{k-1}}{2},\frac{X_k-X_{k-1}}{h}\right)\\ &\quad+\frac{h}{2}F_X\left(a^*_{k-1},\frac{X_k+X_{k-1}}{2},\frac{X_k-X_{k-1}}{h}\right)\\ &\quad-F_{X'}\left(a^*_{k-1},\frac{X_k+X_{k+1}}{2},\frac{X_{k+1}-X_k}{h}\right)\\ &\quad+\frac{h}{2}F_X\left(a^*_{k-1},\frac{X_k+X_{k+1}}{2},\frac{X_{k+1}-X_k}{h}\right)=0 \end{aligned} \tag{16}$$

for $k=1,2,\ldots,N-1$. In the above $a^*_k=(a_k+a_{k+1})/2$ and $X_k=X_h(a_k)$ is the computed value of the solution $X(t)$ at a_k.

We note that (16) is a block tridiagonal system of $M(N-1)$ equations in $M(N-1)$ unknowns which is solved in practice by Newton's method with the accuracy described in Theorem 3, below. We also note that (16) can be used for the initial value problem with $x(a)=x_a$, $x'(a)=x'_a$.

While algorithm (16) is efficient and easy to apply we must also establish that this algorithm satisfies accurate error bounds. This has been done in [9]. These results involve new, mathematical methods obtained primarily from Ref. [7]. The pointwise error estimates, below, have not been obtained for problems of this difficulty. The Richardson error described below, which is almost free, shows that this efficient, easy to apply algorithm is very accurate.

THEOREM 3. *For $h>0$ sufficiently small there exists $C>0$ independent of h so that for any component e of the error $E_h(a_k)=X(a_k)-X_h(a_k)$ we have $|e|\leqslant Ch^2$. In addition, the Richardson solution $X_h^R(t)$, where $X_h^R(a_k)=[4X_{h/2}(a_k)-X_h(a_k)]/3$, has a maximum component, pointwise error satisfying $|e^R|\leqslant Ch^4$ where e^R is any component of $[e_h^R(a_k)=X(a_k)-X_h^R(a_k)$.*

In our specific problem (12) with boundary conditions (3) we have $M=2n+m$ and require transversality conditions at $t=b$ since $X_N=X(b)$ is unknown in (3b) and only the first component is known in (3a). In the former case of (3b) a modification of the argument which resulted in (16) leads to

$$F_X\left(a_{k-1}^*, \frac{X_N+X_{N-1}}{2}, \frac{X_N-X_{N-1}}{h}\right)$$
$$+\frac{h}{2}F_{X'}\left(a_{k-1}^*, \frac{X_N+X_{N-1}}{2}, \frac{X_N-X_{N-1}}{h}\right)=0. \qquad (16)_T$$

Thus, (16) for $k=1, 2, ..., N-1$ and $(16)_T$ provide NM equations in the NM unknowns $X_1, X_2, ..., X_N$. If (3a), then we use the last $m+n$ equations in $(16)_T$ to obtain $NM-n$ equations in the $NM-n$ unknowns.

Finally, we note that simpler, although very meaningful problems, have been solved by the methods described above. These results have appeared as M.S. and Ph.D. theses (see [4, 6]).

IV. THE LINEAR REGULATOR PROBLEMS

We consider the special case where $2f=u^TR(t)u+x^TP(t)x$ in (1), $g=A(t)x+B(t)u$ in (2), and (3) holds where $R(t)$ is positive definite and $P(t)$ is nonnegative definite for all t in $[a, b]$. The "boundary" term $\frac{1}{2}x^T(b)\,Hx(b)$ in (1) for constant nonnegative definite matrix H is easily included in our analysis and we assume for simplicity of exposition that $H=0$.

Using the notation above we define in (14)

$$J_4(X)=\int_a^b F(t, X, X')\,dt, \qquad \text{where}$$

$$f(t, x, u)+\lambda^T g(t, x, u)=\tfrac{1}{2}u^TRu+\tfrac{1}{2}x^TPx+\lambda^T(x'-Ax-Bu),$$
$$F(t, X, X')=\tfrac{1}{2}x_2'R_{x_2'}+\tfrac{1}{2}x_1Px_1+x_3'(x_1'-Ax_1-Bx_2')$$

and

$$L(X)=\frac{1}{2}\int_a^b [X'^T\underline{R}(t)X'+X'^T\underline{\underline{Q}}(t)X+X^T\underline{P}(t)X]\,dt, \qquad (17a)$$

where

$$\underline{R}=\begin{pmatrix}0 & 0 & I\\ 0 & R & -B^T\\ I & -B & 0\end{pmatrix},\quad \underline{Q}=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ -2A & 0 & 0\end{pmatrix},\quad \text{and}\quad \underline{P}=\begin{pmatrix}P & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}. \tag{17b}$$

Finally, we obtain algorithm (16) by noting that

$$F_{X'}=\underline{R}(t)X'+\frac{1}{2}\underline{Q}(t)X=\begin{pmatrix}x_3'\\ Rx_2'-B^Tx_3'\\ x_1'-Bx_2'-Ax_3\end{pmatrix}$$

and (18)

$$F_X=\frac{1}{2}\underline{Q}^T(t)X'+\underline{P}(t)X=\begin{pmatrix}-A^Tx_3+Px_1\\ 0\\ 0\end{pmatrix}.$$

We note that $f_{x_1}=P(t)\,x_1$, $g_{x_1}=A$, $f_{x_2'}=Rx_2'$, $g_{x_2'}=B$ so that using (13) we have

$$F_X=\begin{pmatrix}P(t)\,x_1-A^Tx_3'\\ 0\\ 0\end{pmatrix}\quad\text{and}\quad F_{X'}=\begin{pmatrix}x_3'\\ Rx_2-B^Tx_3'\\ x_1'-Ax_1-Bx_2'\end{pmatrix}$$

which agrees with (18).

V. Equality and Inequality Constraints

The purpose of this section is to indicate that our methods can incorporate equality and inequality constraints. For example, following Hestenes [11, p. 346] we might have the constraints $\varphi_\alpha(t,x,u)\leqslant 0$ $(1\leqslant\alpha\leqslant m')$, $\varphi_\alpha(t,x,u)=0$ $(m'\leqslant\alpha\leqslant\bar{m})$.

Equality constraints are treated as we have done in (12) for the trajectory equation. Multipliers are incorporated as additional components similar to the $x_3'=\lambda$ variable in (12) and their derivations and variations in (15) act linearly. Inequality constraints are treated similarly except that if the constraint equation is inactive (strictly less than zero, below), the multiplier vanishes (see [11, pp. 346–350]). This gives rise to a Kuhn–Tucker condition which will be treated in detail in later work. We will sketch these ideas and also show how to change inequality constraints into equality constraints.

For ease of exposition we now assume $m=2$ and constraint of the form

$$|u| \leqslant K, \tag{19}$$

where K is a positive constant. More general situations can be easily formulated by the methods we now use.

The above constraint is equivalent to

$$\varphi_1 = u - K \leqslant 0 \tag{20a}$$

or

$$\varphi_2 = -u - K \leqslant 0. \tag{20b}$$

In this case, (12) becomes

$$\begin{aligned} F(t, X, X') = f(t, x_1, x_2') + x_3'^T[x_1' - g(t, x_1, x_2')] \\ + x_4'(x_2' - K) + x_5'(-x_2' - K), \end{aligned} \tag{21a}$$

where

$$x_4(t) = \int_a^t \mu_1(s)\, ds, \qquad x_4(a) = 0, \tag{21b}$$

$$x_5(t) = \int_a^t \mu_2(s)\, ds, \qquad x_5(a) = 0 \tag{21c}$$

and

$$X(t) = (x_1^T(t), x_2(t), x_3^T(t), x_4(t), x_5(t)). \tag{21d}$$

In the above we use the multipliers $\mu_\alpha(t)$ for the general constraint $\varphi_\alpha(t, x, u) \leqslant 0$ considered by Hestenes [11, pp. 346–350]. We note that $\varphi_\alpha(t)$ in our case is given in (20). The multipliers $\mu_\alpha(t)$ are piecewise continuous and continuous at each point of continuity of the extremal control $u^*(t)$. Moreover, $\mu_\alpha(t) \geqslant 0$ with $\mu_\alpha(t) = 0$ at each point t for which $\varphi_\alpha(t, x^*(t), u^*(t)) < 0$.

The obvious modifications in (15) and (16) are left to the reader.

Inequality constraints can also be treated by conversion to equality constraints. Thus, for example we can replace φ_1, φ_2 in (20) by introducing the "surplus" variables $x_6(t)$ and $x_7(t)$ with $x_6(a) = x_7(a) = 0$ and

$$\varphi_1 = x'^2 + x_2' - K = 0 \tag{20a$'$}$$

or

$$\varphi_2 = x_7'^2 - x_2' - K = 0. \tag{20b$'$}$$

The multipliers are as given in (21) except that now we have

$$f(t, X, X') = f(t, x_1, x_2') + x_3'([x_1' - g(t, x_1, x_2')] + x_4'(x_6'^2 + x_2' - K) + x_5'(x_7'^2 - x_2' - K). \tag{21b$'$}$$

The advantage of this conversion method over the Kuhn–Tucker methods is that we always deal with equalities. The disadvantage are that more variables are needed (minor) and that originally "linear" problems now are nonlinear.

More complete discussion of inequality constraints and transversality conditions will be given by the third author in other work.

VI. Reduction of the Dimension

The purpose of this section is to indicate that in many instances the dimension of the problem $M = 2n + m$ may be reduced.

The best situation of this type is when $m = n$ and $u(t)$ can be solved as $u(t) = g_1(t, x, x')$. This might happen in the linear regulator problem of Section 4 if $B(t)$ is invertible where the trajectory equation is

$$x'(t) = A(t)\,x(t) + B(t)\,u(t). \tag{22}$$

Often, the linear regulator problem can undergo a partial reduction of dimension. The following results are given in [4].

Let $m < n$, rank $B = m$, and B of the form $[0/B_1]$, where B_1 is $m \times m$. Also, let $p = n - m$ and write

$$A = \left[\frac{A_1}{A_2}\right],$$

where A_1 is $p \times n$ and A_2 is $m \times n$,

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

where x_1 is $p \times 1$ and x_2 is $m \times 1$. Then (22) can be written in the form

$$\begin{aligned} x_1' &= A_1 x \\ x_2' &= A_2 x + B_1 u. \end{aligned} \tag{23}$$

If B is constant and of full rank, then there exists a constant non-singular matrix G such that

$$GB = \left[\frac{0}{B_1}\right]$$

where B_2 is $m \times m$ and non-singular.

If f in (1) is

$$f = \tfrac{1}{2}[u^T R(t) u + x^T P(t) x], \tag{24}$$

where $R(t)$ is positive definite and $P(t)$ is positive semi-definite, then a straightforward calculation leads to a problem in the calculus of variations with (14) of the form

$$F(t, x, x') = \tfrac{1}{2}[x^T P_2 x - 2x_2'^T Q_2 x + x_2'^T R_2 x_2'], \tag{25}$$

where

$$\begin{aligned} P_2 &= A_2^T B_1^{-1^T} R B_1^{-1} A_2 + P, \\ Q_2 &= B_1^{-1^T} R B_1^{-1} A_2, \qquad \text{and} \\ R_2 &= B_1^{-1^T} R B_1^{-1}. \end{aligned}$$

Thus, we have

THEOREM 4. *Under the above hypothesis we have*:

(i) P_2, Q_2 *and* R_2 *are piecewise continuous and bounded on* $[a, b]$,

(ii) $P_2(t)$ *is symmetric and positive semidefinite for all* $t \in [a, b]$,

(iii) $R_2(t)$ *is symmetric and positive definite for all* $t \in [a, b]$,

(iv) $J_2(x) \geq 0$ *for all admissible* x *with equality if and only if* $x(t) \equiv 0$ *for* $t \in [a, b]$, *when* $x_a = 0$,

(v) (x^*, u^*) *is the unique solution to the basic problem* (1)–(3) *if and only if* x^* *is the unique solution to the problem of minimizing* $J_2(x)$ *subject to* (3) *and* $u^* = B_1^{-1}(x_2^{*\prime} - A_2 x^*)$.

The results follow as in the proof of Theorem 1 above and from the obvious calculations (see [4]).

VII. NUMERICAL EXAMPLES

The purpose of this section is to present some numerical examples. Our examples will be for quadratic–linear regulator type problems with simple coefficients since those are the problems where the analytic solutions may be more easily computed. However, in [9] we find nontrivial nonlinear examples with the convergence results contained in Theorem 3. Thus, our examples are simple not because of our methods but because it is difficult to get problems with nice, known solutions for comparison.

We will give two examples but list only one case. In all cases the result of Theorem 3 that the maximum pointwise error of any component of $X(t)$ is $O(h^2)$ holds. More details may be found in [4].

For our first example we have $n = m = 2$,

$$J(x) = \frac{1}{2}\int_0^{\sqrt{2}} \left\{ u^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix} u + x^T \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x \right\} dt,$$

$$x' = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} x + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} u, \qquad \text{and}$$

$$x(0) = \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix}, \qquad x(\sqrt{2}) = \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} e^{-2}.$$

We may solve for $u(t)$ and obtain

$$J_1(x) = \frac{1}{2}\int_0^{\sqrt{2}} \left[x^T \begin{pmatrix} 1 & 0 \\ 0 & 4/3 \end{pmatrix} - 2x^T \begin{pmatrix} 0 & 1/3 \\ 0 & -1 \end{pmatrix} x + x'^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix} x' \right] dt.$$

It can be verified that

$$x(t) = \begin{pmatrix} \sqrt{2}\, e^{-\sqrt{2}t} \\ e^{-\sqrt{2}t} \end{pmatrix} \qquad \text{and} \qquad u(t) = \begin{pmatrix} -3e^{-\sqrt{2}t} \\ (-\sqrt{2}-1)\, e^{-\sqrt{2}} \end{pmatrix}$$

is the optimal solution.

For our second example we have $m = n = 1$; $a = 0$, $b = 1$; $x(0) = x_a = 1$, $x(1)$ is free;

$$J(x) = \frac{1}{2}\int_0^1 (u^2 + 3x^2)\, dt;$$

$$x' = x + B(t)u,$$

where

$$B(t) = \begin{cases} 1 & \text{if } t \text{ in } [0, \tfrac{1}{2}] \\ 0 & \text{if } t \text{ in } (\tfrac{1}{2}, 1]. \end{cases}$$

Pontryagin's maximum principle can be used to establish that $x(1) = e^{3/2}$ and hence

$$x(t) = \begin{cases} e^{2t} & \text{for } t \text{ in } [0, \tfrac{1}{2}] \\ e^{t+1/2} & \text{for } t \text{ in } (\tfrac{1}{2}, 1] \end{cases}$$

$$u(t) = \begin{cases} e^{2t} & \text{for } t \text{ in } [0, \tfrac{1}{2}] \\ 0 & \text{for } t \text{ in } (\tfrac{1}{2}, 1] \end{cases}$$

and

$$P(t) = \begin{cases} e^{-2t} & \text{for } t \text{ in } [0, \tfrac{1}{2}] \\ \tfrac{1}{2}e^{3/2t} - \tfrac{3}{2}e^{1/2+t} & \text{for } t \text{ in } (\tfrac{1}{2}, 1]. \end{cases}$$

The results are obtained exactly as described by (12)–(15) for this example. They are with $h = \frac{1}{8}$,

t_k	$x_1(t)$	Error(h)	Error($h/2$)	Error($h/4$)
0.250	1.6487	2.27 −4	5.66 −5	1.42 −5
0.500	2.7183	1.24 −4	3.09 −5	7.72 −6
0.750	3.4903	3.44 −5	8.55 −6	2.13 −6

t_k	$x_2(t)$	Error(h)	Error($h/2$)	Error($h/4$)
0.250	0.8244	2.52 −4	6.30 −5	1.58 −5
0.500	1.3591	2.10 −4	5.25 −5	1.31 −5
0.750	1.3591	1.05 −4	2.62 −5	6.56 −6

t_k	$x_3(t)$	Error(h)	Error($h/2$)	Error($h/4$)
0.250	4.9018	3.57 −4	8.93 −5	2.23 −5
0.500	5.4366	4.20 −4	1.05 −4	2.62 −5
0.750	6.8667	2.52 −4	6.29 −5	1.57 −5

The reader may observe that the error in each component for $h/2$ is one-fourth the error for h. While we have listed these results the Richardson errors would be much smaller with a ratio of approximately one sixteenth. We note in passing that $x'(\frac{1}{2}-) = 2e$ and that $x'(\frac{1}{2}+) = e$. That is, our extremal solution has a corner at $t = \frac{1}{2}$.

Our final example involves an initial value problem with inequality constraint. Thus, let $m = n = 1$

$$J(x) = \frac{1}{2} \int_0^{\pi} (u^2 - 2x^2)\, dt,$$

$$x' = x + u,$$

and

$$x(0) = 0, \qquad x'(0) = 1.$$

Our constraint is $|u| \leqslant \alpha = \sqrt{6}/2$.

The complete details are contained in [4] but we note that the control $u(t)$ is given by

$$u(t) = \begin{cases} \cos t - \sin t = \sqrt{2} \cos\left(t + \dfrac{\pi}{4}\right) & \text{if } 0 \leqslant t \leqslant \dfrac{7\pi}{12} \\ -\alpha & \text{if } \dfrac{7\pi}{12} < t \leqslant t^* \\ \cos\left(t - t^* \dfrac{7\pi}{12}\right) + \sin\left(t - t^* \dfrac{7\pi}{12}\right) & \text{if } t^* < t \leqslant \pi \end{cases}$$

where $t^* = 7\pi/12 + \ln(-\tan(7\pi/12))$. We also note that our optimal control is free if $0 \leqslant t \leqslant 7\pi/12$ in that $|u| < \alpha$, u is on one boundary if $7\pi/12 < t < t^*$, u is free again in a third interval $t^* < t < t^{**}$ and would be on the other boundary if $t > t^{**}$.

VIII. Abnormal Problems

A complete discussion of abnormal problems is found in [7, pp. 201–223]. Briefly, the characterization of this phenomenon is that the extremal solution is not identically zero but vanishes on nontrivial subintervals $[a', b']$ of $[a, b]$. This phenomenon generalizes the usual concepts of conjugate points or oscillation points where extremal solutions vanish at isolated points of $[a, b]$.

For convenience we consider the following problem. Let $x(t)$ and $u(t)$ be four-vectors. We wish to minimize

$$J(x) = \frac{1}{2}\int_0^\pi (u^T u - x^T x)\, dt$$

subject to

$$x' = B(t)u$$

where (with $e = \pi/4$)

$$B(t) = \begin{cases} \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & 0 \leqslant t \leqslant 2e \\[2ex] \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & 2e < t \leqslant 3e \\[2ex] \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & 3e < t \leqslant \pi \end{cases}$$

We assume boundary conditions or initial value conditions consistent with the extremal solution

$$x(t)=\begin{cases}[2\sin 2t, 2\sin 2t, 0, 0]^T, & 0\leqslant t\leqslant 2e\\ [0, 0, 0, 0]^T, & 2e<t\leqslant 3e\\ [-2\sin(2t-6e), 0, 0, 0]^T, & 3e<t\leqslant \pi.\end{cases}$$

We will see that the nonzero solution to this linear problem vanishes on nontrivial subintervals which is usually impossible because of uniqueness of solutions. The reason for this phenomena is the nature of $B(t)$. Note that in contrast to more usual problems, many values of $u(t)$ lead to the same trajectory since B has a nontrivial nullspace.

For completeness we note that the optimal control $u(t)$ is given by

$$u(t)=\begin{cases}[2\sqrt{2}\cos 2t, 0, 0, 0]^T, & 0\leqslant t\leqslant 2e\\ [0, 0, 0, 0]^T, & 2e<t\leqslant 3e\\ [-2\cos(2t-6e), 0, 0, 0]^T, & 3e<t\leqslant \pi\end{cases}$$

Proceeding as above we set $x_1(t)=x(t)$, $x_2'(t)=u(t)$ and define the multiplier $x_3'(t)$. The above problem is equivalent to

$$\min J_1(x)=\int_0^{4e}\left[\tfrac{1}{2}x_2'^T x_2' - \tfrac{1}{2}x_1^T x_1 + x_3'^T(x_1' - Bx_2')\right] dt.$$

Complete details of computer runs appear in Gibson [6]. However, the error results are essentially the same as described in Theorem 3 and computationally as the examples in the last section.

References

1. M. Athans and P. L. Falb, "Optimal Control: An Introduction to the Theory and Its Applications," McGraw-Hill, New York, 1966.
2. J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, "The Theory of Splines and their Applications," Mathematics in Science and Engineering Series, Volume 38, New York, 1967.
3. W. E. Bosarge and O. G. Johnson, Error bounds of high order accuracy for the state regulator problem via piecewise polynomial approximation, *SIAM J. Control* **9** (1971), 15–28.
4. R. Dalpatadu, "New Numerical Methods in Optimal Control Theory," Ph.D. Dissertation, Southern Illinois University, May 1986.
5. R. Dalpatadu and J. Gregory, Methods for optimal control problems with bilinear objective functions and linear trajectories, *in* "Conference Proceedings on Matrix Methods, Auburn, Alabama, March 1986," to appear.
6. C. Gibson, "Numerical Methods for Abnormal Quadratic Form Problems in Optimal Control Theory," Master's Thesis, Southern Illinois University at Carbondale, April 1986.

7. J. GREGORY, "Quadratic Form Theory and Differential Equations," Academic Press, New York, 1980.
8. J. GREGORY, M. ZEMAN, AND M. BADIEY, A finite element approximation for the initial-value problem for nonlinear second order differential equations. II. Systems, *J. Math. Anal. Appl.* **115** (1986), 131–139.
9. J. GREGORY AND R. S. WANG, A complete numerical solution for the m-dependent variable, nonlinear extreme problem in the calculus of variations, submitted for publication.
10. P. HENRICI, "Discrete Variable Methods in Ordinary Differential Equations," Wiley, New York, 1962.
11. M. R. HESTENES, "Calculus of Variations and Optimal Control Theory," Wiley, New York, 1966.
12. M. R. HESTENES, Multiplier and gradient methods, *Optim. Theory Appl.* **4** (1967), 303–320.
13. M. D. INTRILIGATOR, "Mathematical Optimization and Economic Theory," Prentice–Hall, Englewood Cliffs, NJ, 1971.
14. H. B. KELLER, "Numerical Methods for Two-point Boundary Value Problems," Ginn (Blaisdell), Boston, 1968.
15. D. E. KIRK, "Optimal Control Theory," Electrical Engineering Series, Prentice–Hall, Englewood Cliffs, NJ, 1970.